Information Sciences 422 (2018) 204–217

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Local attribute reductions for decision tables

Guilong Liu*, Zheng Hua, Jiyang Zou

School of Information Science, Beijing Language and Culture University, Beijing 100083, China

ARTICLE INFO

Article history:
Received 30 September 2016
Revised 29 August 2017
Accepted 4 September 2017
Available online 6 September 2017

Keywords:
Attribute reduction
Decision table
Discernibility matrix
Equivalence relation
Reduction algorithm

ABSTRACT

Attribute reduction is among the most important areas of research in rough sets. This paper investigates the types of local attribute reduction for decision tables. We propose the concepts of lth decision class lower approximation reduction, lth decision class reduction, and lth decision class β-reduction for decision tables, and provide their corresponding reduction algorithms via discernibility matrices. We also establish the relationship between positive-region reduction and the lth decision class β-reduction, and report a case study using the University of California–Irvine dataset to verify the theoretical results.

1. Introduction

Proposed by Pawlak [23,24], rough set theory is a useful tool to study decision systems characterized by insufficient and incomplete information. In more than 30 years of development, rough set theory has been successfully applied to many fields, such as artificial intelligence, machine learning, pattern recognition, decision analysis, and knowledge discovery in databases. A basic notion supporting the application of rough sets is the concept of attribute reduction [5,6,8,11,22,26,37]. The aim of attribute reduction in the context of a decision table is to obtain a minimal subset of attributes with the same classification power as the entire set of attributes.

Many significant results pertaining to attribute reduction have been obtained in the literature [1–3,8–10,12,13,15,16,18–20,25,33,34], and a variety of attribute reductions have been defined with different criteria. For example, positive-region reduction, also introduced by Pawlak [24], keeps the positive region unchanged. Wu [35] considered attribute reduction in decision tables based on the Dempster–Shafer theory of evidence. Ziarko [38] introduced the concept of β-reduction based on variable precision rough sets. This type of reduction is an important generalization of classical reduction that preserves the sum of objects in β lower approximations of all decision classes. Slezak [30] constructed a framework for approximate Bayesian networks and provided examples related to rough set-based attribute reduction. Slezak and Ziarko [29] proposed a Bayesian rough set model suitable for the identification and elimination of redundant attributes. Mi et al. [20] proposed the concepts of β lower distribution reduction and β upper distribution reduction based on variable precision rough sets. To expedite the reduction, Yang et al. [36] found that minimal elements in a discernibility matrix are sufficient to find all reductions. Inuiguchi et al. [9] proposed a variable precision dominance-based rough set reduction approach. Many new attribute reduction methods have recently been developed as well. Jia et al. [11] summarized 22 definitions of attribute reduction and proposed a generalized attribute reduction that considers the data and user preferences.

* Corresponding author.
E-mail address: liuguilong@blcu.edu.cn (G. Liu).

http://dx.doi.org/10.1016/j.ins.2017.09.007

This paper can be viewed as a refinement of typical attribute reduction. The aim is to obtain local decision rule acquisition algorithms for decision tables. Positive-region reduction can acquire all rules in a decision table but requires a large amount of time and space. There are many decision classes for a decision table. However, we are sometimes interested in the characterization of a special decision class in the hope that a certain property of this class remains constant. That is, we expect that an approach can acquire rules for some decision classes but does not compute simultaneously for all classes. For example, we hope that the lower approximation for some decision classes remains unchanged. To address this issue, we propose the concepts of lth decision class lower approximation reduction and lth decision class reduction. As an extension of such types of reduction, we define the concept of the lth decision class β-reduction for decision tables. All these types of reduction can be viewed as local reduction for decision tables. We study the relationship between positive-region reduction and lth decision class β-reduction, and show that the lth decision class lower approximation reduction algorithm can find all positive-region reductions. This provides a new and simple algorithm for positive-region reduction.

Skowron and Rauszer [28,31] were the first to propose the concept of the discernibility matrix and report important results [28]. Discernibility matrix-based attribute reduction is now a commonly used and efficient method. This paper defines discernibility matrices for different types of reduction.

The remainder of this paper is structured as follows: In Section 2, we review some basic notions related to variable precision rough sets and their extensions. Section 3 proposes the concept of lth decision class lower approximation reduction for decision tables and formulates a corresponding algorithm. Section 4 studies lth decision class reduction for decision tables and provides a reduction algorithm for it. As an extension of lth decision class reduction, Section 5 considers lth decision class β-reduction and its algorithm. Section 6 discusses the relationship between positive-region reductions and lth decision class β-reductions, and Section 7 reports a case study to verify our theoretical results. Section 8 contains the conclusions of this study.

2. Variable precision rough sets and their extensions

In this section, we review variable precision rough sets and their extensions. Let $U = \{x_1, x_2, \ldots, x_n\}$ be a finite set of objects called the universal set and $P(U)$ be the power set of U. Suppose that R is a binary relation on U. Recall that the left and right R-relative sets of an element x in U are defined as

$$l_R(x) = \{y | y \in U, yRx\} \text{ and } r_R(x) = \{y | y \in U, xRy\},$$

respectively. R is referred to as serial if $r_R(x) \neq \emptyset$ for each $x \in U$. R is referred to as reflexive if, for each $x \in U$, $x \in r_R(x)$; R is referred to as symmetric if, for each $x \in U$, $r_R(x) = l_R(x)$, and is called transitive if, for each $x, y, z \in U$, $y \in r_R(x)$ and $z \in r_R(y)$ imply $z \in r_R(x)$. R is referred to as an equivalence relation if R is reflexive, symmetric, and transitive.

Recall that if X is a subset of a universal set U, the characteristic function λ_X of X is defined, for each $x \in U$, as follows:

$$\lambda_X(x) = \begin{cases} 1, & x \in X \\ 0, & x \notin X \end{cases}.$$

We need some lemmas.

Lemma 2.1. *Let A and B be subsets of U. Then, $\lambda_A = \lambda_B$ if and only if $A = B$.*

Similarly, for any given binary relation R on U, for $x, y \in U$, $\lambda_R(x, y)$ is equal to one if xRy, and is zero otherwise.

For a serial relation R on U, suppose that $M_R = (a_{ij})_{n\times n}$ is the relational matrix of R, i.e., $a_{ij} = \lambda_R(x_i, x_j)$. We define $n \times n$ matrices as follows: $N_R = \begin{pmatrix} \frac{1}{|r_R(x_1)|} & & & \\ & \frac{1}{|r_R(x_2)|} & & \\ & & \ddots & \\ & & & \frac{1}{|r_R(x_n)|} \end{pmatrix}$, *and*

$$W_R = N_R M_R = \begin{pmatrix} \frac{a_{11}}{|r_R(x_1)|} & \frac{a_{12}}{|r_R(x_1)|} & \cdots & \frac{a_{1n}}{|r_R(x_1)|} \\ \frac{a_{21}}{|r_R(x_2)|} & \frac{a_{22}}{|r_R(x_2)|} & \cdots & \frac{a_{2n}}{|r_R(x_2)|} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{a_{n1}}{|r_R(x_n)|} & \frac{a_{n2}}{|r_R(x_n)|} & \cdots & \frac{a_{nn}}{|r_R(x_n)|} \end{pmatrix}.$$

Lemma 2.2. *Let R and S be serial relations on U. Then, $W_R = W_S$ if and only if $R = S$.*

Proof. We must show that $x_iRx_j \Leftrightarrow x_iSx_j$. $x_iRx_j \Leftrightarrow \lambda_R(x_i, x_j) = 1$. If x_iRx_j, by $W_R = W_S$, $\frac{1}{|r_R(x_i)|} = \frac{\lambda_R(x_i, x_j)}{|r_R(x_i)|} = \frac{\lambda_S(x_i, x_j)}{|r_S(x_i)|}$, which means that x_iSx_j. Similarly, we can show that x_iSx_j implies x_iRx_j. □